# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| **1. REPORT DATE** *(DD-MM-YYYY)* | **2. REPORT TYPE** | **3. DATES COVERED** *(From - To)* |
|---|---|---|
| 18-11-2013 | R&D Status Report (Quarterly) | 17-08-2013 to 16-11-2013 |

**4. TITLE AND SUBTITLE**
Construction of a Bacterial Cell that Contains Only the Set Of Essential Genes Necessary to Impart Life

**5a. CONTRACT NUMBER**
HR0011-12-C-0063

**5b. GRANT NUMBER**
NA

**5c. PROGRAM ELEMENT NUMBER**
NA

**6. AUTHOR(S)**
John Glass

Tony Yee

**5d. PROJECT NUMBER**
NA

**5e. TASK NUMBER**
NA

**5f. WORK UNIT NUMBER**
NA

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
J. Craig Venter Institute
9704 Medical Center Drive
Rockville, MD 20850

**8. PERFORMING ORGANIZATION REPORT NUMBER**
HR0011-12-C-0063.6

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
DARPA, MTO
Prog: Living Foundries ATGC
675 N Randolph St
Arlington, VA 22203

**10. SPONSOR/MONITOR'S ACRONYM(S)**
DARPA

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**
NA

**12. DISTRIBUTION / AVAILABILITY STATEMENT**

Approved for public release; distribution is unlimited

**13. SUPPLEMENTARY NOTES**
NA

**14. ABSTRACT**

We have made significant progress on all fronts of the project. On the Top Down approach to genome minimization, we have used the Essential (E), Non-essential (N), and Impaired (I) gene categories to make steady progress with gene and gene cluster deletions. To date, we have removed approximately 273 kb from the Mycoplasma mycoides JCVI-syn1.0 genome. The resultant 806 kb genome is viable and grows with a normal growth rate.

Results on the Bottom Up approach are substantial. We have continued work on the reduced genome design (RGD1) based on the N/E/I categorization system and completed testing on 6 out of 8 segments from the RGD1 design. All six segments are viable in a 7/8th wild type background. In separate experiments, we have also tested genomes containing 2 (Segments 2+6) or 3 (Segments 2+6+8) minimized segments. These two genome are also viable.

The effort to modularize the genome is in progress. A tRNA module was constructed and inserted into the M. mycoides syn1.0 genome. The resulting genome is viable following transplantation.

Preliminary work aimed at genome complementation has continued. Experiments using a transposon based approach have been successful, and allowed recovery of growth rate on a genome with an impaired growth rate.

**15. SUBJECT TERMS**
NA

| **16. SECURITY CLASSIFICATION OF:** | | | **17. LIMITATION OF ABSTRACT** | **18. NUMBER OF PAGES** | **19a. NAME OF RESPONSIBLE PERSON** |
|---|---|---|---|---|---|
| **a. REPORT** | **b. ABSTRACT** | **c. THIS PAGE** | | | Tony Yee |
| UU | UU | UU | UU | 11 | **19b. TELEPHONE NUMBER** *(include area code)* 301 795 7133 |

# *Construction of a Bacterial Cell that Contains Only the Set of Essential Genes Necessary to Impart Life*

| | |
|---|---|
| Report Title: | R&D Status Report (Quarterly) |
| Report Number: | HR0011-12-C-0063.6 |
| Reporting Period: | August 17, 2013 to November 16, 2013 |
| Contract No.: | HR0011-12-C-0063 |
| Performing Organization: | J. Craig Venter Institute<br>9704 Medical Center Drive<br>Rockville, MD 20850<br>USA |
| Author of Report: | John Glass, Anthony Yee |
| Principal Investigators: | Hamilton Smith, Clyde Hutchison |


**Abstract**

We have made significant progress on all fronts of the project. On the Top Down approach to genome minimization, we have used the Essential (E), Non-essential (N), and Impaired (I) gene categories to make steady progress with gene and gene cluster deletions. To date, we have removed approximately 273 kb from the *Mycoplasma mycoides* JCVI-syn1.0 genome. The resultant 806 kb genome is viable and grows with a normal growth rate.

Results on the Bottom Up approach are substantial. We have continued work on the reduced genome design (RGD1) based on the N/E/I categorization system and completed testing on 6 out of 8 segments from the RGD1 design. All six segments are viable in a 7/8$^{th}$ wild type background. In separate experiments, we have also tested genomes containing 2 (Segments 2+6) or 3 (Segments 2+6+8) minimized segments. These two genome are also viable.

The effort to modularize the genome is in progress. A tRNA module was constructed and inserted into the *M. mycoides* syn1.0 genome. The resulting genome is viable following transplantation.

Preliminary work aimed at genome complementation has continued. Experiments using a transposon based approach have been successful, and allowed recovery of growth rate on a genome with an impaired growth rate.



Sponsored by
Defense Advanced Research Projects Agency
Microsystems Technology Office (MTO)
Program: Living Foundries: Advanced Tools and Capabilities for Generalizable Platforms (ATCG)
Issued by DARPA/CMO under Contract No: HR0011-12-C-0063

Table of Contents:



## Summary

The goal of the project is to create a cell that contains only the set of genes that are essential for life under ideal laboratory conditions. We are working to minimize *Mycoplasma mycoides* JCVI-syn1.0 (the synthetic version of *Mycoplasma mycoides* subsp *capri*) using two approaches:

- Top Down: remove genes and clusters of genes one (or a few) at a time, proceeding only if the reduced strain is viable, with a reasonable growth rate
    - We previously reported the results of a transposon study was conducted and allowed us to categorize genes as Essential (E), Non-essential (N), or Impaired (I)
        - This categorization scheme has been the basis of most of our subsequent work
    - The *M. mycoides* genome has been reduced to 806 kb using the Tandem Repeat Endonuclease Cleavage (TREC) strategy
    - By making gene cluster deletions in the 1/8th genome segments and assembling the segments into a genome, a 790 kb genome was made
        - The genome was viable, but had a doubling time of 82 minutes
- Bottom Up: design a reduced genome based on our best Tn5 gene disruption and deletion data (RGD1), and synthesize it
    - Synthesis from oligonucleotides
        - The first six 1/8$^{th}$ genome molecules (Segments 1, 2, 4, 6, 7 and 8) have been tested and found to be viable
        - Testing on Segments 3 and 5 is in progress
        - Genomes containing reduced Segments 2+6 and Segments 2+6+8 have been tested and found to be viable.
        - Assembly of a genome containing Segments 1+2+4+6+7+8 is in progress

The initial modularization experiments are progressing. A 5.3 kb tRNA module containing the 30 tRNA genes plus the necessary promoters and terminators was constructed and sequence verified. The module was inserted into the genome and found to be viable.

Work has continued on the genome complementation front. We now have a working transposon based approach.

## Introduction

The goal of this research project is to build a minimal bacterial cell that contains only the genes that are required for life in ideal laboratory conditions. The pursuit of a minimized cell is critical to the advancement of biology, both as a pathway for understanding the basic requirements for replication and as a chassis for creating an optimized platform for any number of possible applications.

We previously reported that the *Mycoplasma mycoides* JCVI-syn1.0 genome was successfully reduced from 1078 kb to 779 kb; however, while the 779 kb genome was viable, the growth rate was far too slow to allow follow up experiments at an acceptable pace. Using the N, E, I gene categories, the genome has been reduced to 806 kb, but with a normal doubling time.

Synthesis of a newly designed reduced genome (RGD1, 539 kb) was completed at SGI. All 6 of the 1/8$^{th}$ genome segments tested so far have proven to be viable. Testing of the final two pieces will be completed in November 2013. We have no reason to believe that these two segments will not be viable.

A preliminary tRNA module was designed, constructed, and introduced into the *M. mycoides* syn1.0 genome and found to be viable.

Experiments on a transposon-based genome complementation system are being conducted.

**Methods, Assumptions and Procedures**

TOP DOWN APPROACH

The plan here was to start with the full size 1078 kb *M. mycoides* JCVI-syn1.0 synthetic genome. The smallest viable genome to date was 779 kb, created using an 8-piece strategy to build genomes containing multiple deletions. However, the transplanted genome proved to grow very poorly in liquid culture. Using a less-reduced but faster growing strain as a starting point, we have successfully created a 790 kb strain, with an 82 min doubling time.

We have continued to use the TREC strategy to make iterative deletions in the mycoplasma genome. Targeting the N category genes and clusters is proving to be effective (further discussed in the Results and Discussion section). We have made a series of strains that are progressively reduced with little to no reduction in growth rate.

BOTTOM UP APPROACH

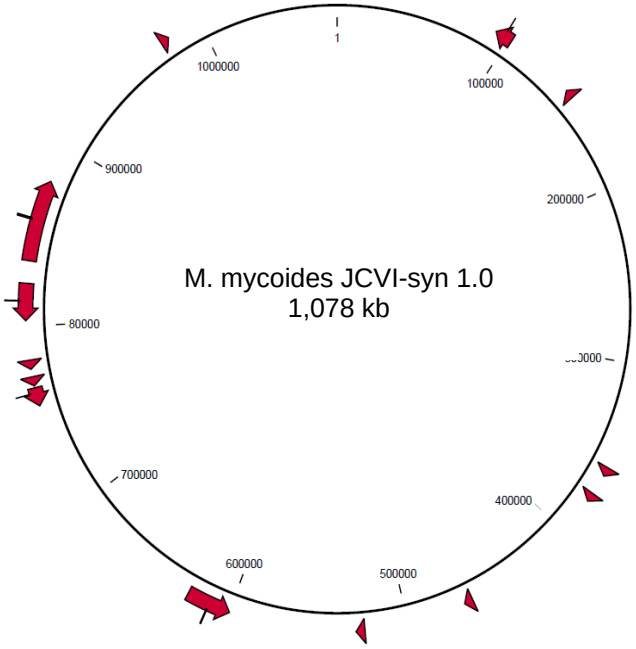
Synthesis from oligonucleotides: A new eight piece genome design was completed using the N, E, I gene classification system (RGD1, 539 kb). The designs were provided to SGI, and all 8 segments have been sequence verified and returned. Each segment is being inserted into a 7/8 normal landing pad strain for testing. So far, 6 of the segments have been found to support cellular life. Testing on the remaining two will be complete in the next two weeks.

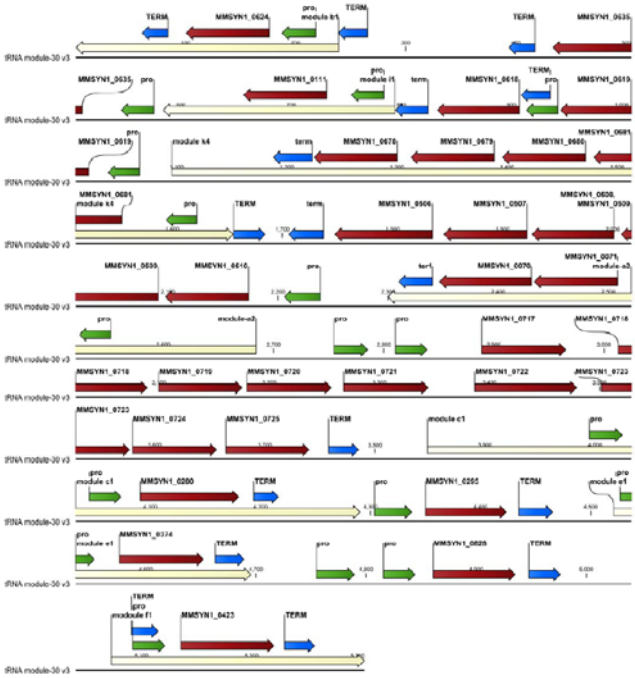
We have begun testing combinations of reduced genome segments. Two genomes have been assembled containing Segments 2+6 or Segments 2+6+8. These combinations are viable.

MODULARIZATION

To test gene modularization, we have organized the 30 tRNA genes of *M. mycoides* into a single contiguous module. The module contains the coding regions, as well as the promoters and terminators needed for regulation. The tRNA genes are naturally distributed around the genome in 13 loci.

**Figure 1**



***(a) Natural distribution of tRNA genes in M. mycoides.*** *The tRNA gene clusters have been enlarged in Fig.1(a) to show the direction of transcription. The M. mycoides JCVI-syn1.0 genome has 8 single tRNA genes and 5 clusters of 2 to 9 genes, for a total of 30.*

***(b) tRNA module design.*** *The 30 tRNA genes have been relocated into a single module. (Green arrows represent promoters. Red arrows show tRNA genes. Blue arrows are terminators.)*

Each of the 13 loci was synthesized by PCR using syn1.0 as the template, cloned in E. coli and then joined together into a single cassette with appropriate yeast markers. Sequencing revealed 3 differences from the design, but none occur in the tRNA genes themselves. The cassette was inserted into syn1.0 to replace the largest cluster of 9 tRNAs at 10 o'clock on the genome map. The resulting genome is viable after transplantation. We will use the Green Monster Method developed by Yo Suzuki to disrupt the 12 other tRNA loci around the genome.

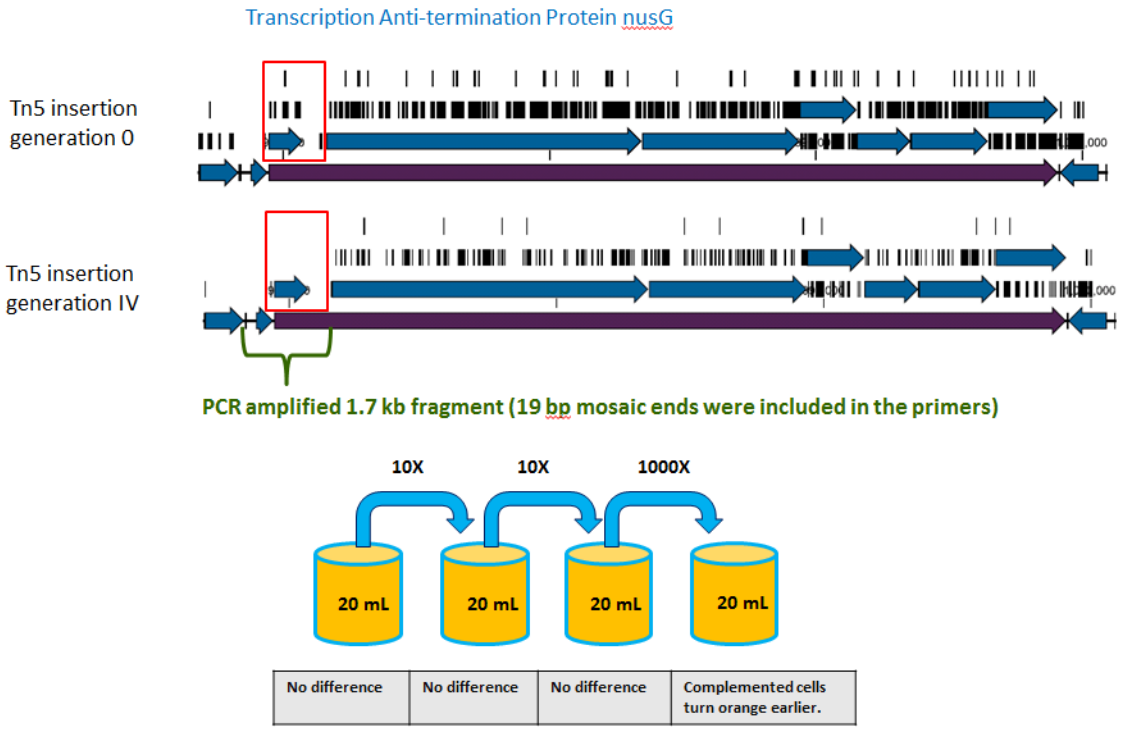
GENOME COMPLIMENTATION

We have begun initial experiments aimed at enabling genetic complementation to restore desirable phenotypes to deleted strains, for example, RGD1. A system capable of quickly adding deleted genes back into a genome would be a powerful tool to help de-convolute growth-retarding synthetic effects.

Currently, we are testing a strategy based on a transposon approach. An impaired growth deletion (transcription anti-termination protein nusG) was PCR amplified using primers that contain the 19 bp mosaic ends from the Tn5 transposon construct. The PCR product was transformed into the impaired genome in the presence of a transposase, and the growth rate was recovered.

This approach has several advantages. It is much more efficient from an effort and time perspective than using a plasmid based system. It is reasonable to expect that we can use this approach in a library fashion to examine the complementation effects of several genes at a time. A schematic of the initial approach is shown below in Figure 2:

**Figure 2**



## Results and Discussion

MODULARIZATION

We have moved ahead with modularization of the tRNA loci. Construction was difficult because the complex secondary structures of the tRNAs interfered with the synthesis process. The module is complete, sequenced, and inserted into the genome. The genome containing the tRNA module is viable, and we will proceed to disrupt the loci around the genome using the Green Monster Method.

TOP DOWN APPROACH

Iterative deletions using the TREC based approach are making steady progress toward a minimal genome. A table outlining the progress to date is shown below. Strains D13 has been designed, but testing for viability has not been completed.

**Table 1**

| Strains | DT(min) | Genome Size (bp) | # of Genes Deleted |
|---|---|---|---|
| **syn1.0** | 64 | 1,078,809 | 0 |
| **syn1.0D6 RE** | | 1,062,183 | 17 |
| **DISs** | | 1,048,690 | 31 |
| **D1** | | 979,083 | 68 |
| **D2** | | 969,069 | 74 |
| **D3** | | 944,159 | 90 |
| **D4** | | 931,710 | 97 |
| **D5** | | 923,647 | 102 |
| **D6** | 67 | 908,931 | 108 |
| **D7** | | 877,942 | 135 |
| **D8** | | 866,271 | 155 |
| **D9** | 64 | 844,265 | 173 |
| **D10** | 65 | 828,901 | 181 |
| **D11** | | 816,807 | 194 |
| **D12** | | 805,506 | 201 |
| **D13*[1]** | | ~795,000 | 209 |
| **D14*[2]** | | ~784,000 | 214 |

*[1] *In progress; assembled (transplantation data is not yet available)*
*[2] *Planned*

Sequential deletion of genes and clusters of genes is not the way that we ultimately expect to pursue a minimal genome design. While we are making steady progress in this approach, the greater outcome of performing the top down deletions is the generation of information regarding unanticipated interactions between elements of the genome that are difficult or impossible to identify using the Bottom Up approach. De-convolution of synthetic effects will be difficult work, which we hope to delay as long as possible with information gained from the top down approach.

BOTTOM UP APPROACH

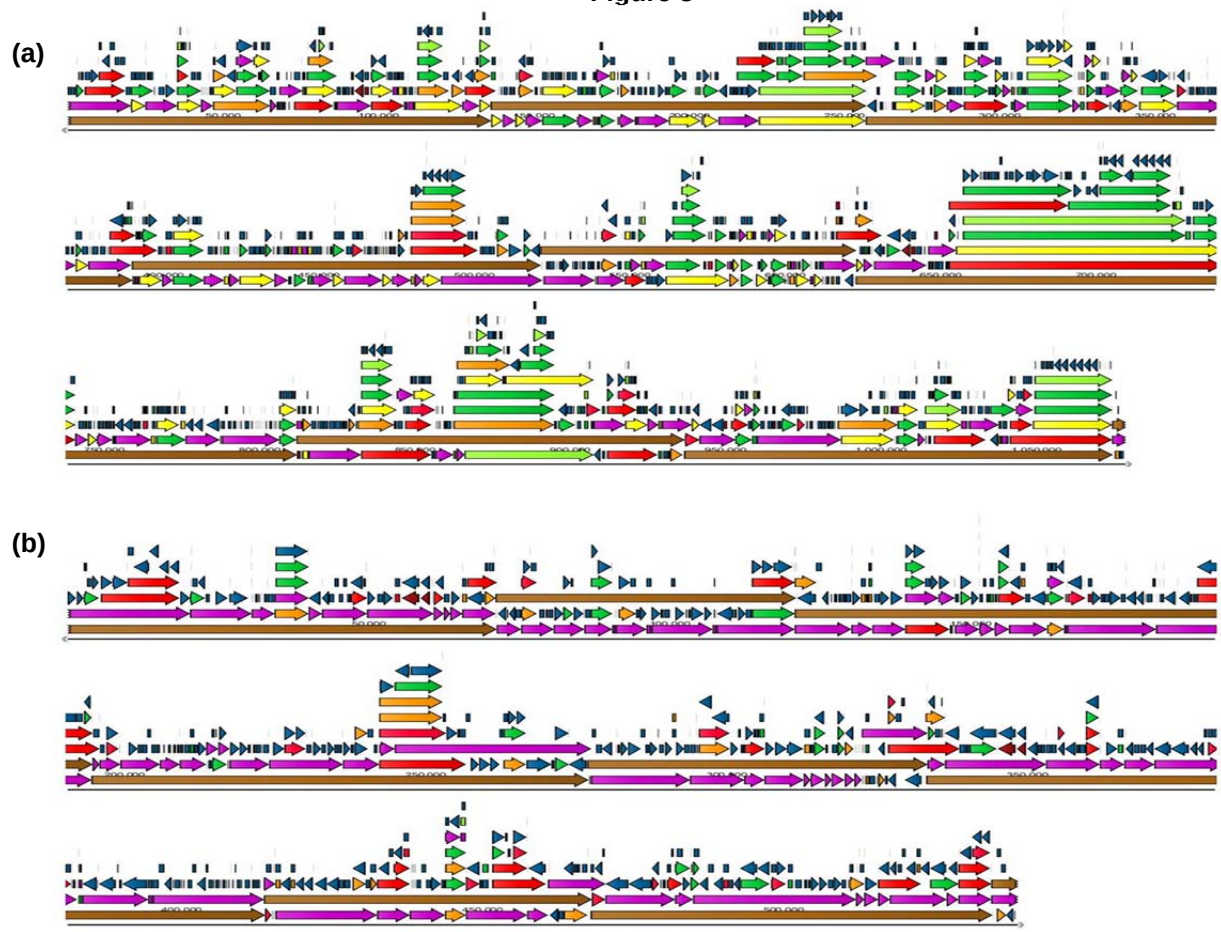
A new genome design (RGD1) was completed using the N, E, I classification system. The design essentially involves removing all of the N (non-essential) genes and then checking to see that we have not disturbed the promoters and terminators necessary for expression of the remaining E (essential), and I (impaired growth if deleted) genes.

Figure 3, below, shows the contrast between (a) the *M. mycoides* JCVI-syn1.0 genome and (b) the new genome design.

**Figure 3**

**(a)**

**(b)**




***(a) The 1,078 kb M. mycoides JCVI-syn1.0 genome, showing the genes retained in the new genome design (purple arrows)****. Note that genes in the new design are interspersed around the genome, comingled with genes that were removed from the design;* ***(b) The new design of a minimized, 539 kb M. mycoides genome****. Contrast the spacing between the retained genes (purple) with the spacing in (a).*

*If a lethal deletion (red) falls entirely within a region deleted in the design (yellow) then there is an inconsistency between the design and the deletion data. Where the data from different deletions disagrees, green trumps red because failure to get a transplant is a negative result that could result from some problem with that particular transplantation reaction.*

The table below shows the length of each 1/8th genome Not I segment in the *M. Mycoides* JCVI-syn1.0 genome, and the newly designed RGD1 genome. The new design represents a 50% reduction in total genome length.

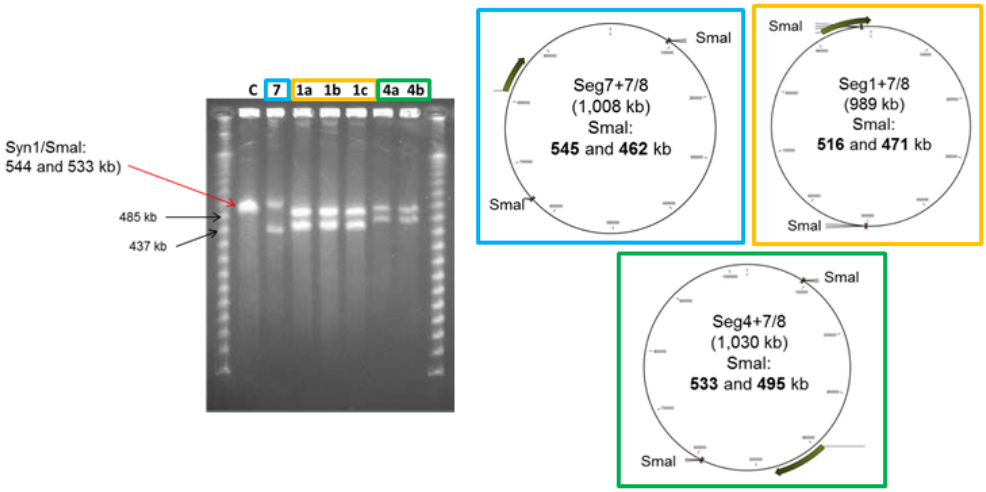
**Table 2**

| Not I Fragment # | *M. Mycoides* JCVI-syn1.0 Length (bp) | RGD1 Designed Length (bp) | (RGD1)/(*M. Mycoides* JCVI-syn1.0) |
|---|---|---|---|
| 1 | 140,739 | 75,732 | 0.54 |
| 2 | 120,912 | 49,888 | 0.41 |
| 3 | 133,208 | 73,958 | 0.56 |
| 4 | 131,623 | 82,531 | 0.63 |
| 5 | 101,708 | 56,501 | 0.56 |
| 6 | 189,357 | 80,747 | 0.43 |
| 7 | 124,976 | 54,482 | 0.44 |
| 8 | 137,887 | 66,717 | 0.48 |
| Total | 1,080,410 | 540,566 | |
| Overlaps | -1,601 | -1,601 | |
| Genome Length | 1,078,809 | 538,955 | 0.50 |

In this reporting period, we have made a great deal of exciting progress on the RGD1 genome. We have received all 8 of the designed segments from SGI DNA, and the results of testing shown below:

Segment 1: Swapping into a 7/8th background and transplantation is complete. Initial results show that the transplant is viable and supports cellular growth. Results confirmed by PCR and CHEF gel analysis. The CHEF gel analysis is shown below, in Figure 4.

**Figure 4**




CHEF gel analysis of RGD Segment 1, Segment 4 and Segment 7. Each segment was swapped into a 7/8 wild type background and transplanted. The results show that the viable transplants produce the expected bands following restriction with SmaI.

Segment 2: Testing was completed prior to this reporting period. The minimized segment supports cellular growth at an acceptable growth rate in a 7/8$^{th}$ wild type background.

Segment 3: The segment was recently received from SGI DNA. Assembly into a genome and transplantation are ongoing. Preliminary results will be available by the end of November.

Segment 4: Swapping into a 7/8$^{th}$ background and transplantation is complete. Initial results show that the transplant is viable and supports cellular growth. Results confirmed by PCR and CHEF gel analysis (see Fig. 4, above).

Segment 5: The segment was recently received from SGI DNA. Assembly into a genome and transplantation are ongoing. Preliminary results will be available by the end of November.

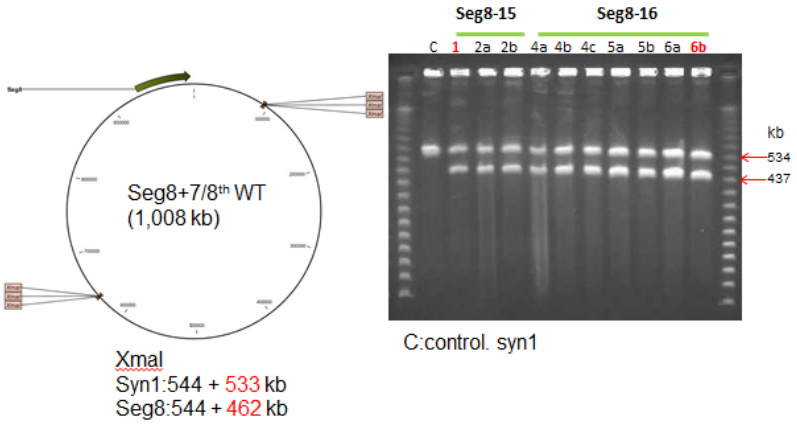
Segment 6: The initial test of segment 6 indicated that while viable, the growth rate was very slow following transplantation.

Interestingly, subsequent cultures prepared for growth rate testing then recovered and grew at approximately normal rates. The deletion of one non-essential (N) region caused the joining of two essential (E) genes. A transcription terminator hairpin downstream of one of the E genes was influencing the other. Mutation in Segment 6 sequence reduced the stability of the transcription terminator hairpin (C=>T at position 672,318 of Segment6+7/8; position 766,571 of syn1.0) allowed the recovered growth rate.

Segment 7: Swapping into a 7/8$^{th}$ background and transplantation is complete. Initial results show that the transplant is viable and supports cellular growth. Results confirmed by PCR and CHEF gel analysis(see Fig. 4, above).

Segment 8: Testing has been completed. The minimized segment supports cellular growth at an acceptable growth rate in a 7/8$^{th}$ wild type background. The CHEF gel analysis is shown below:

**Figure 5**




CHEF gel analysis of RGD Segment 8 in a 7/8 wild type background following transplantation. The expected bands following restriction with XmaI were produced.

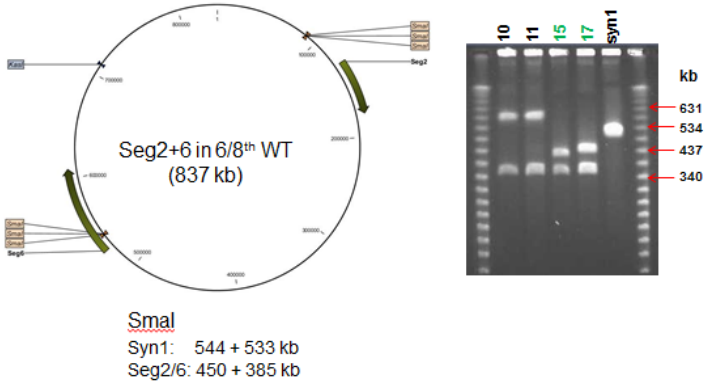
In summary, six segments of RGD1 have been tested through transplantation. ***All six segments are viable*****.**

In addition to testing the individual segments in a wild type background, we have begun testing combinations of viable, minimized segments. The genomes are produced via 8-piece assembly in

yeast, and transplanted. In contrast, individual segments are tested by swapping into a 7/8$^{th}$ landing pad strain.

Segments 2+6 in a 6/8$^{th}$ wild type background: We began the effort to combine RGD1 Segments 2 and 6 into one genome as soon as the results indicated that the individual segments were viable. The results are very encouraging – the combination of Segments 2 and 6 is viable. The transplant has been confirmed by PCR and CHEF gel analysis.

**Figure 6**




CHEF gel analysis of RGD Segments 2+8 in a 6/8 wild type background following transplantation. The expected bands following restriction with SmaI were produced.

Segments 2+6+8 in a 5/8$^{th}$ wild type background: ***Preliminary results indicate that the combination of Segments 2+6+8 is viable and displays a reasonable growth rate***. The colonies have been genotyped by PCR. CHEF gel analysis is not yet complete.

## Conclusions

Tasks from the Statement of Work for Year 1:

**Task 1:** **Complete a detailed global Tn5 transposon mutagenesis insertion map.**
The Tn5 transposon insertion map was submitted with the initial quarterly report.

Due: Month 6; Status - complete

**Task 2:** **Delete up to 27 large gene clusters**
We have reduced the genome size of M. mycoides JCVI-syn1.0 from 1079 kb to 779 kb through the deletion of some 30 clusters, representing a ~30% reduction. We are moving forward with an 848 kb genome because it grows at a higher rate than the smaller 779 kb version.

Due: Month 12; Status – complete

**Task 3:** **Construct a preliminary modular map of the genome**
The design of a modular map of the genome is complete was presented in the May 2012 quarterly report

Due: Month 12; Status – complete

**Task 4: Make new transposon insertion map. Identify non-essential small 2-4 gene clusters. Delete small clusters.**

A new transposon study was performed and previously reported. A table showing the resulting N/E/I categorization system was presented in May 2013 (Deliverable 1). We continue to work toward a minimal cell chassis using multiple simultaneous approaches.

Due: Month 18 (November 2013); Status – in progress. We propose delaying the completion of this task until the testing of the RGD1 1/8$^{th}$ segments is complete.

We have no reason to believe that the remaining two 1/8$^{th}$ segments will not be viable. If this assumption holds true, by the end of 2013 we will know if the complete, 50% reduced RGD1 genome is viable.

It would be the first living organism that is radically different from any existing cell, and whose genome has been entirely computationally designed and chemically synthesized. It will have the smallest genome of any known independently growing cell on earth. We will then use Tn5 mutagenesis to discover if the classifications of any of the remaining genes have changed and to determine and report which additional genes might be removed (Deliverable 2).

## Planned Activities for the Next Reporting Period

1. Continue the Top Down minimization of our synthetic genome (near term) and conduct a new transposon study on the minimized genome (within approximately 4-6 months)
2. Complete testing of the 8 segments of the RGD1 genome as part of our Bottom Up strategy.
3. If the remaining 2 RGD segments are viable, construct a genome containing all 8 segments and test its viability (we expect this could be completed by the end of 2013)
4. Continue with verification and testing of the tRNA gene module.
5. Continue development of genome complementation strategies

## Program Financial Status

| In Process & Completed Tasks | Planned Expend | Actual Expend (Cumulative to Date) | % Budget Completion | At Completion | Latest Revised Estimate | Remarks |
|---|---|---|---|---|---|---|
| Task 1 | $305,646 | $305,646 | 100% | $305,646 | $305,646 | Completed |
| Task 2 | $826,256 | $798,351 | 97% | $798,351 | $826,256 | Completed |
| Task 3 | $43,487 | $43,487 | 100% | $43,487 | $43,487 | Completed |
| Task 4 | $634,981 | $441,396 | 70% | N/A | $634,981 | N/A |
| Cumulative | $1,810,370 | $1,588,880 | 88% | N/A | $1,810,370 | N/A |

There is no management reserve or unallocated resources. The financial data presented is current through Sept. 2013.

Based on the currently authorized work:

- Is current funding sufficient for the current fiscal year? Yes
- What is the next fiscal year funding requirement at current anticipated levels? $1,214,151.00
- Have you included in the report narrative any explanation of the above data and are they cross-referenced? Not applicable; current funding is sufficient for the current fiscal year.